Homework Assignment #2

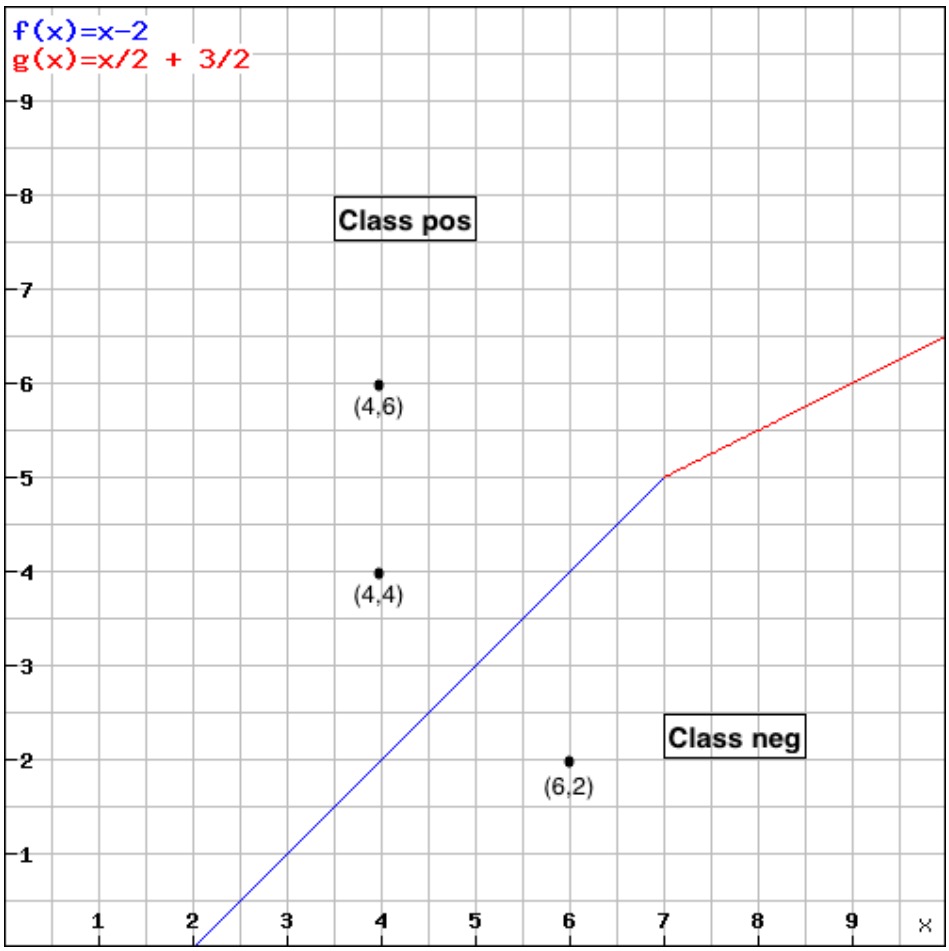
Part 1

Question 1:

Consider using single nearest-neighbor for a task with two continuous features that both have the range [0, 10]. Suppose you are given the following three training instances. Show the decision regions (their boundaries and associated class labels) induced by these three instances when Euclidean distance is used to identify the nearest neighbors for test instances.

x_1	x_2	y
4	4	pos
4	6	pos
6	2	neg

Solution :



Question 2:

Using the k-d tree and the training set displayed in the figure below, show how the nearest neighbor for *x(q) = (7, 10)* would be found. For each step in the search, show the distance to the current node, the best distance found so far, the best node found so far, and the contents of the priority queue. You should use Euclidean distance and assume that the coordinates of the instances in the figure below are at integer values.

Coordinates of Instances:

a: (2,11)
b: (3,12)
c: (5,10)
d: (2, 8)
e: (2, 4)
f: (6, 3)
g: (9, 1)
h: (12, 5)
i: (10,10)
j: (13,12)

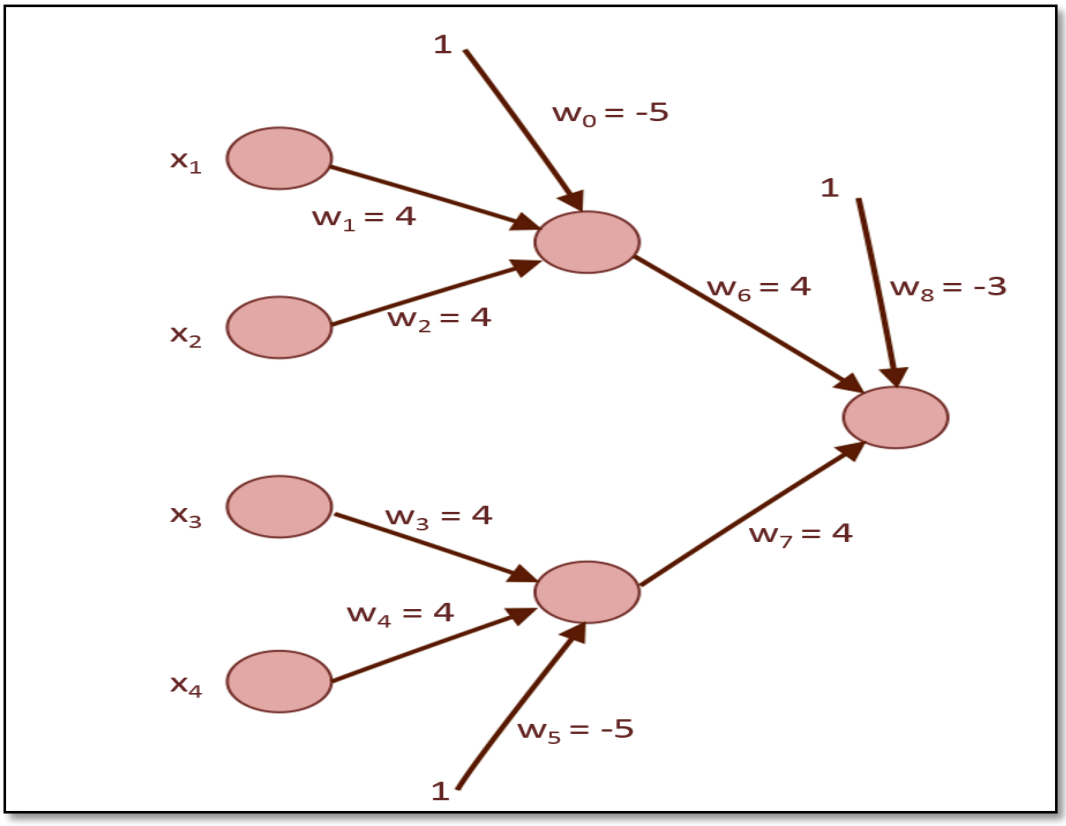
Solution :

	distance	best distance	best node	priority queue
		∞		(f,0)
pop f	$\sqrt{50} = 7.071068$	$\sqrt{50} = 7.071068$	f	(h,0) (c,1)
pop h	$\sqrt{50} = 7.071068$	$\sqrt{50} = 7.071068$	f	(i,0) (c,1) (g,5)
pop i	3	3	i	(c,1) (j,3) (g,5)
pop c	2	2	c	(e,0) (b,0) (j,3) (g,5)
pop e	$\sqrt{61} = 7.81025$	2	c	(b,0) (d,0) (j,3) (g,5)
pop b	$\sqrt{20} = 4.47214$	2	c	(d,0) (j,3) (a,4) (g,5)
pop d	$\sqrt{29} = 5.38516$	2	c	(j,3) (a,4) (g,5)
pop j				

bound ≥ best distance
3 ≥ 2
return c which is the best node

Question 3:

Show a neural network that represents the logical function $y = (x1 \wedge x2) \vee (x3 \wedge x4)$. Specifically, show the network topology, weights and biases. You should assume that hidden and output units use sigmoid output functions, and an output-unit activation of 0.5 or greater represents a *true* prediction for *y*.



Question 4:

Given the neural network below, calculate and show the weight changes that would be made to the two weights shown in red ($w_{D,A}$ and $w_{F,D}$) by one step of the online version of backpropagation for the training instance shown. Assume that the hidden and output units use sigmoid functions, the network is being trained to minimize squared error, and the learning rate is 0.1 (with no momentum term). Blue weights in the figure indicate bias parameters.

Solution :

net$^{(d)}$ = w0 + $\sum_{i=1}^{n} w_i x_i$

If j is output unit -> $\delta j = Oj(1 - Oj)(y_j - Oj)$ and $\Delta w_{ji} = \eta \delta j Oi$

If j is hidden unit -> $\delta j = Oj(1 - Oj) \sum_k \delta_k w_{kj}$ and $\Delta w_{ji} = \eta \delta j Oi$

net$^{(c)}$ = 1*0.5 + 0*1 + 1*-0.5 = -1

net$^{(d)}$ = 0*-0.5 + 1*1 + 1*-1 = 0

O_1^2 = 1/(1 + $e^{-(-1)}$) = 0.268941

O_2^2 = 1/(1 + e^{-0}) = 0.5

O_1^3 = 1/(1+$e^{-(1*0 + -2*0.268941 + 1*0.5)}$) = 0.490531

O_2^3 = 1/(1+$e^{-(-2*0.268941 + -2*0.5 + 1*-0.5)}$) = 0.11528257

δ_1^3 = 0.490531(1-0.490531)(1-0.490531) = 0.127322

δ_2^3 = 0.11528257(1-0.11528257)(0-0.11528257) = -0.011758

Weight changes to $w_{F,D}$ be represented by Δw_{22}^2

Weight changes to $w_{D,A}$ be represented by Δw_{21}^1

Δw_{22}^2 = 0.1 * δ_2^3 * O_2^2

= 0.1 * - 0.011758 * 0.5

= - 0.0005879

δ_2^2 = O_2^2 (1 - O_2^2) [$\delta_1^3 * w_{12}^2 + \delta_2^3 * w_{22}^2$]

= 0.5 * (1-0.5) * [0.127322 * 1 + -0.011758 * -2]

= 0.037709

Δw_{21}^1 = 0.1 * δ_2^2 * O_1^1

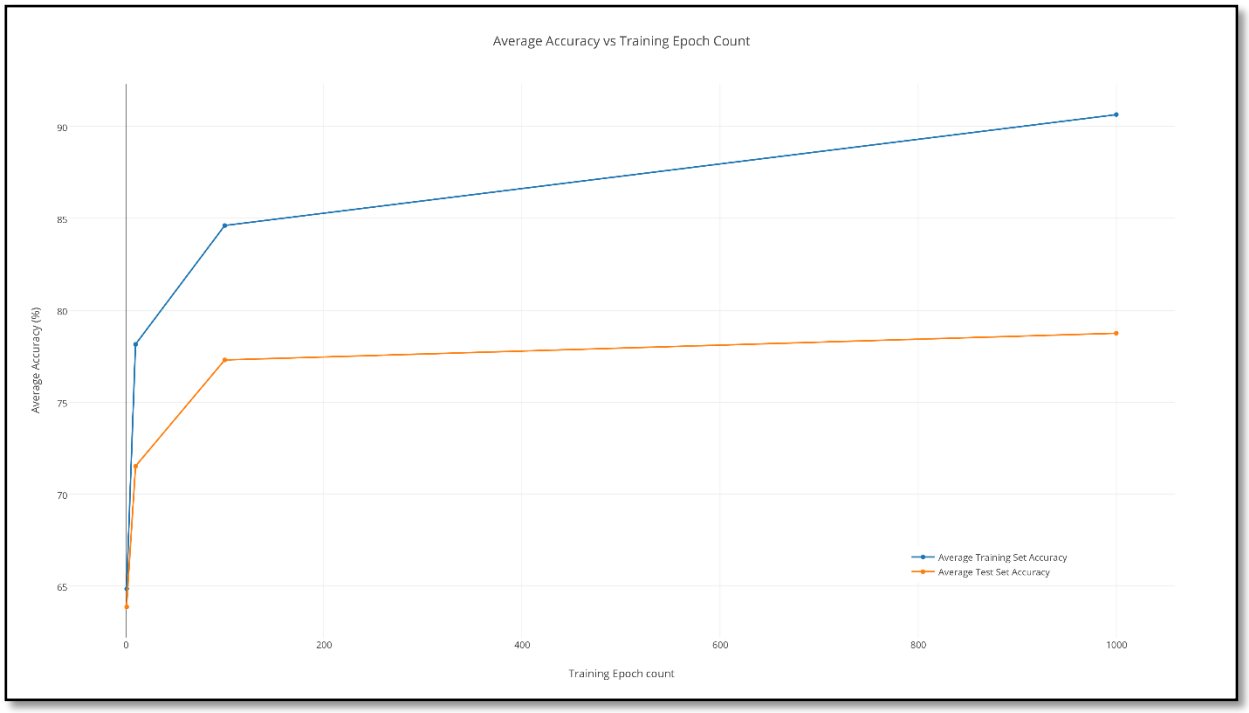
= 0,1 * 0.037709 * 0

= 0

Part 2

Question 5:

Using stratified 10-fold cross validation, make a set of graphs showing the average training- and test-set accuracy after 1, 10, 100, and 1000 training epochs (i.e. passes through the entire training set). When the activation on the output unit (i.e. the value computed by the sigmoid) > 0.5, you should treat this as a prediction for the 'Mine" class.



Question 6:

Plot an ROC curve for a run of stratified 10-fold cross validation with a learning rate of 0.1 and 100 training epochs. You should pool the classifications from the 10 test sets to make one curve. Use the activation of the output unit as the measure of confidence that a given test instance is positive. You should consider 'Mine' to be the positive class.